

Research Article

The Maximal Strichartz Family of Gaussian Distributions: Fisher Information, Index of Dispersion, and Stochastic Ordering

Alessandro Selvitella

Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S 4L8

Correspondence should be addressed to Alessandro Selvitella; aselvite@math.mcmaster.ca

Received 26 April 2016; Accepted 7 August 2016

Academic Editor: Timothy R. Marchant

Copyright © 2016 Alessandro Selvitella. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We define and study several properties of what we call *Maximal Strichartz Family of Gaussian Distributions*. This is a subfamily of the family of Gaussian Distributions that arises naturally in the context of the *Linear Schrödinger Equation* and Harmonic Analysis, as the set of maximizers of certain norms introduced by Strichartz. From a statistical perspective, this family carries with itself some extrastructure with respect to the general family of Gaussian Distributions. In this paper, we analyse this extrastructure in several ways. We first compute the *Fisher Information Matrix* of the family, then introduce some measures of *statistical dispersion*, and, finally, introduce a *Partial Stochastic Order* on the family. Moreover, we indicate how these tools can be used to distinguish between distributions which belong to the family and distributions which do not. We show also that all our results are in accordance with the dispersive PDE nature of the family.

1. Introduction

The most important multivariate distribution is the *Multivariate Normal (MVN) Distribution*. To fix the notation, we give here its definition.

Definition 1. One says that a random variable X is distributed as a Multivariate Normal Distribution if its probability density function (pdf)

$$f_X : \mathbf{R}^n \longrightarrow \mathbf{R} \quad (1)$$

takes the form

$$f_X(x_1, \dots, x_n) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right), \quad (2)$$

where $\mu := E[X] \in \mathbf{R}^n$ is the mean value vector and $\Sigma := \text{Var}(X) \in \text{Sym}_n^+$ is the $n \times n$ positive definite symmetric Variance-Covariance Matrix.

Its importance derives mainly (but not only) from the *Multivariate Central Limit Theorem* which has the following statement.

Theorem 2. Suppose that $X = (x_1, \dots, x_n)^T$ is a random vector with Variance-Covariance Matrix Σ . Assume also that $E[x_i^2] < +\infty$ for every $i = 1, \dots, n$. If X_1, X_2, X_3, \dots is a sequence of i.i.d. random variables distributed as X , then

$$\frac{1}{n^{1/2}} \sum_{i=1}^n (X_i - E[X_i]) \longrightarrow^d \text{MVN}(0, \Sigma), \quad (3)$$

where \rightarrow^d represents the convergence in distribution.

Due to its importance, several authors have tried to give characterizations of this family of distributions. See, for example, [1, 2] for an extended discussion on multivariate distributions and their properties. Here, we concentrate on characterizing the MVN through variational principles, such as the maximization of certain functionals. A well-known characterization of the Gaussian Distribution is through the maximization of the *Differential Entropy*, under the constraint of

fixed variance Σ . We focus on the case of when the support of the pdf is the whole Euclidean Space \mathbf{R}^n .

Theorem 3. *Let X be a random variable whose pdf is f_X . The Differential Entropy $h(X)$ is defined by the following functional:*

$$h(X) := - \int_{\mathbf{R}^n} f_X(x) \log f_X(x) dx. \tag{4}$$

The Multivariate Normal Distribution has the largest Differential Entropy $h(X)$ amongst all the random variables X with equal variance Σ . Moreover, the maximal value of the Differential Entropy $h(X)$ is $h(\text{MVN}(\Sigma)) = (1/2) \log[(2\pi e)^n |\Sigma|]$.

We refer to Appendix for a proof of this well-known theorem. This characterization is, in some sense, not completely satisfactory because it is given just with the restriction of fixed variance. A more general characterization of the *Gaussian Distribution* has been given in a setting which, at first sight, seems very far, and it is the one of Harmonic Analysis and Partial Differential Equations. We first introduce the so-called *admissible exponents*.

Definition 4. Fix $n \geq 1$. One calls a set of exponents (q, r) admissible if $2 \leq q, r \leq +\infty$ and

$$\frac{2}{q} + \frac{n}{r} = \frac{n}{2}. \tag{5}$$

Remark 5. These exponents are characteristic quantities of certain norms, the Strichartz Norms, naturally arising in the context of Dispersive Equations and can vary from equation to equation. We refer to [3] for more details.

Here is the precise characterization of the Multivariate Normal Distribution, through Strichartz Estimates.

Theorem 6 (see [4–7]). *Suppose $n = 1$ or $n = 2$. Then, for every (q, r) and (\tilde{q}, \tilde{r}) admissible and for every $u_0 \in L^2_x(\mathbf{R}^n)$ such that $\|u_0\|_{L^2(\mathbf{R}^n)}^2 = 1$, we have*

$$\|e^{-it\Delta} u_0\|_{L^q_t L^r_x} \leq S(n, q, r), \tag{6}$$

where $S_h(n, q, r) = S_h(n, r)$ is the Sharp Homogeneous Strichartz Constant, defined by

$$S_h(n, r) := \sup \left\{ \|u\|_{L^q_t L^r_x(\mathbf{R} \times \mathbf{R}^n)} : \|u\|_{L^2_x(\mathbf{R}^n)}^2 = 1 \right\} \tag{7}$$

and given by

$$S_h(n, r) = 2^{n/4 - n(r-2)/2r} r^{-n/2r}. \tag{8}$$

Moreover, inequality (6) becomes an equality if and only if $|u_0|^2$ is the pdf of a Multivariate Normal Distribution.

For several other important results on Strichartz Estimates, we refer to [8–11] and the references therein.

Remark 7. This characterization does not need the restriction of fixed variance as the one achieved using the Differential Entropy Functional and so it is, in some sense, more “general.” The result is conjectured to be true for any dimension $n \geq 1$. See, for example, [7], where the optimal constant has been computed in any dimension $n \geq 1$, under the hypothesis that the maximizers are Gaussians also in dimension $n \geq 3$.

We refer to [7] for the relation of this result with harmonic analysis and restriction theorems.

Strichartz Estimates are a fundamental tool in the problem of global well-posedness of PDEs and measure a particular type of dispersion (see, e.g., [3–5, 7, 12, 13]). Strichartz Estimates bring with themselves some interesting statistical features and this is what we want to analyse in the present paper.

The symmetries of the functional in (6) give rise to a family of distributions that we call *Maximal Strichartz Family of Gaussian Distributions*:

$$\mathcal{F} := \left\{ p(t, x) = \left(\frac{\pi}{2}\right)^{-n/2} |R^T R|^{-1/2} |\lambda^2 + 16t^2|^{-n/2} \right. \tag{9}$$

$$\left. \cdot e^{-2(x-x_0-vt)^T (R^T R)^{-1} (x-x_0-vt) / (\lambda^2 + 16t^2)} \right\};$$

$$(t, \lambda) \in \mathbf{R} \times \mathbf{R}, (x_0, v) \in \mathbf{R}^n \times \mathbf{R}^n, R \in SO(n) \}. \tag{10}$$

We refer to Section 2 for its precise construction. This is a subfamily of the family of Gaussian Distributions and, among the other things, it has the feature that the *Mean Vector* μ and the *Variance-Covariance Matrix* Σ depend on common parameters. Therefore, from a statistical perspective, this family carries with itself some extrastructure with respect to the general family of Gaussian Distributions. This extrastructure becomes evident from the form of the *Fisher Information Metric* of the family.

Theorem 8. *Consider $p(t, x)$, a probability distribution function belonging to the Maximal Strichartz Family of Gaussian Distributions \mathcal{F} , defined in (9). The vector of parameters θ , indexing \mathcal{F} , is given by*

$$\theta := \left(x_0^T, v_0^T, \lambda, (R^T R)_{ij}, t \right)^T. \tag{11}$$

Then, the Fisher Information Matrix of $p(t, x)$ is given

(i) *in the spherical case ($R^T R = \sigma^2 Id$) by*

$$I(\theta) = \begin{pmatrix} \frac{1}{\sigma^2} Id & \frac{t}{\sigma^2} Id & 0 & 0 & \frac{1}{\sigma^2} v_0 \\ \frac{t}{\sigma^2} Id & \frac{t^2}{\sigma^2} Id & 0 & 0 & \frac{t}{\sigma^2} v_0 \\ 0 & 0 & \frac{\lambda^2 n}{8} & n \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2) & 2\lambda t n \\ 0 & 0 & n \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2) & \frac{n}{32} \frac{(\lambda^2 + 16t^2)^2}{\sigma^4} & \frac{nt}{\sigma^2} (\lambda^2 + 16t^2) \\ \frac{1}{\sigma^2} v_0 & \frac{t}{\sigma^2} v_0 & 2\lambda t n & \frac{nt}{\sigma^2} (\lambda^2 + 16t^2) & \frac{|v_0|^2}{\sigma^2} + 32nt^2 \end{pmatrix}; \tag{12}$$

(ii) in the elliptical case ($R^T R = \sigma_i^2 Id$; see Section 3 for the precise definition) by

$$I(\theta) = \begin{pmatrix} \frac{1}{\sigma_i^2} Id & \frac{t}{\sigma_i^2} Id & 0 & 0 & \frac{1}{\sigma_i^2} v_0^j \\ \frac{t}{\sigma_i^2} Id & \frac{t^2}{\sigma_i^2} Id & 0 & 0 & \frac{t}{\sigma_i^2} v_0^j \\ 0 & 0 & \frac{\lambda^2 n}{8} & \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2) & 2\lambda t n \\ 0 & 0 & \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2) & \frac{1}{32} \frac{(\lambda^2 + 16t^2)^2}{\sigma^4} & \frac{t}{\sigma^2} (\lambda^2 + 16t^2) \\ \frac{1}{\sigma_i^2} v_0^j & \frac{t}{\sigma_i^2} v_0^j & 2\lambda t n & \frac{t}{\sigma^2} (\lambda^2 + 16t^2) & \sum_{i=1}^n \frac{|v_0^j|^2}{\sigma_i^2} + 32nt^2 \end{pmatrix}. \tag{13}$$

Remark 9. Technically, the only possible case inside the Maximal Strichartz Family of Gaussian Distributions is when $R^T R = Id_{n \times n}$, since $R \in SO(n)$ (the spherical case, with $\sigma^2 = 1$). The form of the Fisher Information Matrix, in that case, is simplified to a lower dimension. Nevertheless, the computation performed in the way we did gives the possibility to compute a distance (in some sense centred at the Maximal Strichartz Family of Gaussian Distributions) between members of the Maximal Strichartz Family of Gaussian Distributions and other Gaussian Distributions, for which the orthogonal matrix condition $R^T R = Id_{n \times n}$ is not necessarily satisfied. In particular, it can distinguish between Gaussians evolving through the PDE flow (see Section 2) and Gaussians which do not.

Remark 10. We believe that using the flow of a Partial Differential Equation is a natural way to produce probability density functions, in particular, in this case, since the flow of the PDE, that we are using, preserves the probability constraint. See Section 2.2 for more details on this comment.

As we said, Strichartz Estimates are a way to measure the dispersion caused by the flow of the PDE to which they are related. In statistics, dispersion explains how stretched or squeezed is a distribution. A measure of statistical dispersion is a nonnegative real number which is small for data which are very concentrated and increases as the data become more spread-out. Common examples of measures of statistical dispersion are the variance, the standard deviation, the range, and many others. Here, we connect the two closely related concepts (dispersion in statistics and PDEs) by introducing some measures of statistical dispersion like the *Index of Dispersion* in Definition 38 (see Section 4) which reflect the dispersive PDE nature of the *Maximal Strichartz Family of Gaussian Distributions*.

Definition 11. Consider the norms $\|\cdot\|_a$ and $\|\cdot\|_b$ on the space of Variance-Covariance Matrices Σ and $\|\cdot\|_c$ on the space of mean values μ . One defines the following *Index of Dispersion*:

$$\mathcal{F}_M^{abc} := \|\Sigma(0)\|_a \times \frac{\|\Sigma(t)\|_b}{\|\mu(t)\|_c^4}, \tag{14}$$

with $t \neq 0$ and where $\mu(t)$ is as follows:

$$\mu(t) := x_0 + vt, \tag{15}$$

while $\Sigma(t)$ is given by

$$\Sigma(t) := \frac{1}{4} (\lambda^2 + 16t^2) R^T R. \tag{16}$$

One calls \mathcal{F}_M^{abc} the *abc*-Dispersion Index of the Maximal Family of Gaussians and one calls

$$\mathcal{F}_M^a := \|\Sigma(0)\|_a \tag{17}$$

a-Static Dispersion Index of the Maximal Family of Gaussians.

We compute this *Index of Dispersion* for our family of distributions and show that it is consistent with PDE results. We refer to Definition 38 for more details.

Another important concept in probability and statistics is the one of *Stochastic Order*. A *Stochastic Order* is a way to consistently put a set of random variables in a sequence. Most of the *Stochastic Orders* are partial orders, in the sense that an order between the random variables exists, but not all the random variables can be put in the same sequence. Many different *Stochastic Orders* exist and have different applications. For more details on *Stochastic Orders*, we refer to [14]. Here, we use our *Index of Dispersion* to define a *Stochastic Order* on the *Maximal Strichartz Family of Gaussian Distributions* and see how there are natural ways of partially ranking the distributions of the family (see Section 5), in agreement with the flow of the PDE.

Definition 12. Consider two random variables X_1 and X_2 such that $\mu_{X_1}(\theta_1) = \mu_{X_2}(\theta)$, for any θ_1 and θ_2 . One says that the two random variables are ordered according to their Dispersion Index \mathcal{F} if and only if the following condition is satisfied:

$$\begin{aligned} X_1 < X_2 &\iff \\ \mathcal{F}(X_2) &\leq \mathcal{F}(X_1). \end{aligned} \tag{18}$$

Remark 13. In this definition the index \mathcal{F} can vary according to the context and the choices of the norms in the definition of the index.

An important tool which will be fundamental in our analysis is what we call *1/α-Characteristic Function* (see Section 2 and [7, 15]). We conclude the paper with an appendix in which, among the other things, we use the concept of *1/α-Characteristic Function* to define generalized types of *Momenta* that exist also for the *Multivariate Cauchy Distributions*.

2. Construction of the Maximal Strichartz Family of Gaussian Distributions

This section is devoted to the construction of the *Maximal Strichartz Family of Gaussian Distributions*; see Figure 1. This

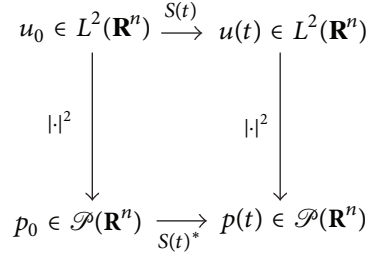


FIGURE 1

is basically done through PDE methods. The program is the following.

- (1) We define *1/α-Characteristic Functions*.
- (2) We prove that if u_0 generates a probability distribution $p_0(x)$, then $u(t, x) = e^{it\Delta} u_0$ (see below its precise definition) still defines a probability distribution $p_t(x) = |u(t, x)|^2$.
- (3) By means of *1/α-Characteristic Functions*, we give the explicit expression of $u(t, x)$ the generator of the family.
- (4) We use symmetries and invariances to build the complete family \mathcal{F} .

2.1. The 1/α-Characteristic Functions. Following the program, we first need to introduce the tool of *1/α-Characteristic Functions* to characterize \mathcal{F} . It is basically the Fourier Transform, but, differently from the *Characteristic Function*, the average is not taken with the pdf, but with a power of the pdf.

Definition 14. Consider $u : \mathbf{R}^n \rightarrow \mathbf{C}$ to be a Schwartz function, namely, a function belonging to the space

$$\mathcal{S}(\mathbf{R}^n) := \{f \in C^\infty(\mathbf{R}^n) \mid \|f\|_{\alpha,\beta} < \infty \forall \alpha, \beta\}, \tag{19}$$

with α and β being multi-indices, endowed with the following norm:

$$\|f\|_{\alpha,\beta} = \sup_{x \in \mathbf{R}^n} |x^\alpha D^\beta f(x)|. \tag{20}$$

Moreover, suppose that

$$\int_{\mathbf{R}^n} |u|^\alpha = 1; \tag{21}$$

namely, $|u|^\alpha$ defines a continuous probability distribution function. Then, one defines

$$\phi_\alpha^u(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-ix \cdot \xi} u(x) dx. \tag{22}$$

One calls $\phi_\alpha^u(\xi)$ the *1/α-Characteristic Function* of u . Moreover, one defines the *Inverse 1/α-Characteristic Function* by

$$\psi_\alpha^u(x) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{ix \cdot \xi} \phi_\alpha^u(\xi) d\xi. \tag{23}$$

We refer to the note [15] for examples and properties of $1/\alpha$ -Characteristic Functions and to Appendix for a simple straightforward application of this tool. In particular, we notice that $\psi_\alpha^u(x) = u(x)$.

Remark 15. If u is complex valued (not just real valued) and, for example, $\alpha = n \in \mathbf{N}$, then there are n -distinct complex roots of $|u|^2$. In our discussion, this will not create to us any problem, because our process starts with u and produces $|u|^2$. We want to remark that the map $|u|^\alpha \mapsto u$ is a multivalued function. For this reason, we cannot reconstruct uniquely a generator, given the family that it generates. See formula (39) below and [15] for more details.

Remark 16. We could define $1/\alpha$ -Characteristic Functions for more general functions $u : X \rightarrow F$ with X a locally compact abelian group and F a general field. We do not pursue this direction here and we will leave it for a future work. We notice that $\phi_\alpha^u(\xi)$ can be considered also as a $1/\alpha$ -Expected Value:

$$E_{|u|^2}^{1/\alpha} [e^{-ix\xi}] = \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-ix\xi} u(x) dx. \tag{24}$$

2.2. Conservation of Mass and Flow on the Space of Probability Measures. In this subsection, we show that if $p_0(x) = |u_0|^2$ defines a probability distribution, then also $p_t(x) = |e^{it\Delta}u_0|^2$ defines a probability distribution. This is mainly a consequence of the property of $e^{it\Delta}$ of being a unitary operator.

Theorem 17. Consider $\mathcal{P}(\mathbf{R}^n)$, the set of all probability distributions on \mathbf{R}^n and $u : (0, \infty) \times \mathbf{R}^n \rightarrow \mathbf{C}$ a solution to (27). Then u induces a flow in the space of probability distributions.

Proof. Consider $u_0 : \mathbf{R}^n \rightarrow \mathbf{C}$ such that $\|u\|_{L^2(\mathbf{R}^n)} = 1$; so $p_0(x) := |u_0(x)|^2$ is a probability distribution on \mathbf{R}^n . Consider $u(t, x)$, the solution of (27) with initial datum u_0 . Then

$$\begin{aligned} \partial_t \int_{\mathbf{R}^n} |u|^2 &= \partial_t \int_{\mathbf{R}^n} u\bar{u} = \int_{\mathbf{R}^n} \partial_t (u\bar{u}) \\ &= \int_{\mathbf{R}^n} (\partial_t u\bar{u} + u\partial_t \bar{u}) \\ &= \int_{\mathbf{R}^n} \Re \left[\bar{u} \left(\frac{i}{2} \Delta u \right) - u \left(\frac{i}{2} \Delta \bar{u} \right) \right] = 0. \end{aligned} \tag{25}$$

So $\partial_t \int_{\mathbf{R}^n} |u|^2 = 0$ and hence

$$\int_{\mathbf{R}^n} |u(t, x)|^2 = \int_{\mathbf{R}^n} |u_0(x)|^2 = 1. \tag{26}$$

Therefore, for every $t \in \mathbf{R}$, $p(t, x) := |u(t, x)|^2$ is a probability distribution. \square

Remark 18. This situation is in striking contrast with respect to the heat equation, where if you start with a probability distribution as initial datum, instantaneously the constraint of being a probability measure is broken.

2.3. Fundamental Solution for the Linear Schrödinger Equation Using $1/\alpha$ -Characteristic Functions. In this subsection, we solve the Linear Schrödinger Equation

$$i\partial_t u(t, x) = \Delta u(t, x), \quad (t, x) \in (0, \infty) \times \mathbf{R}^n, \tag{27}$$

with initial datum $u_0(x) = e^{-|x|^2} \in \mathcal{S}(\mathbf{R}^n)$. This will produce a natural generator $u(t, x)$ of a family of probability distributions, due to Theorem 17.

We first notice that the initial datum becomes a probability density function, if and only if multiplied by a constant. But, since the equation is linear, we can do everything without that constant and then include it at the end.

Remark 19. These computations are well known, but we perform them in detail here, in order to clarify what we will compute in the context of $1/\alpha$ -Characteristic Functions.

Since $u_0(x) \in \mathcal{S}(\mathbf{R}^n)$, then also $\partial_t u(t, x) \in \mathcal{S}(\mathbf{R}^n)$ and $\Delta u(t, x) \in \mathcal{S}(\mathbf{R}^n)$. So, we can apply the $1/2$ -Characteristic Function to both sides of (27) and get

$$i\partial_t \phi_2^u = -|\xi|^2 \phi_2^u, \tag{28}$$

whose solution is

$$\phi_2^u(\xi, t) = e^{i|\xi|^2 t} \phi_2^u(\xi, 0). \tag{29}$$

We now need to compute the $1/\alpha$ -Characteristic Function of the initial datum and then the Inverse $1/2$ -Characteristic Function of $\phi_2^u(t, \xi)$ to get the explicit form of the solution. We have

$$\begin{aligned} \phi_2^u(0, \xi) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-ix\xi} u(0, x) dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-ix\xi} e^{-|x|^2} dx \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-(|x|^2 + ix\xi - |\xi|^2/4)} e^{-|\xi|^2/4} dx \\ &= \frac{e^{-|\xi|^2/4}}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-|x - i\xi/2|^2} dx \end{aligned} \tag{30}$$

by using contour integrals. We notice that, with a simple change of variables, we have

$$\begin{aligned} 2^{n/2} \int_{\mathbf{R}^n} e^{-|x - i\xi/2|^2} dx &= 2^{n/2} \int_{\mathbf{R}^n} e^{-|x|^2} dx \\ &= \int_{\mathbf{R}^n} e^{-|x|^2/2} dx = (2\pi)^{n/2}. \end{aligned} \tag{31}$$

Hence

$$\phi_2^u(0, \xi) = \frac{e^{-|\xi|^2/4}}{(2\pi)^{n/2}} \pi^{n/2} = \frac{e^{-|\xi|^2/4}}{2^{n/2}}. \tag{32}$$

With this, we can conclude

$$\begin{aligned}
 u(t, x) &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbf{R}^n} e^{i|\xi|^2 t + ix \cdot \xi} \frac{e^{-|\xi|^2/4}}{2^{n/2}} = \frac{1}{2^n} \frac{1}{\pi^{n/2}} \\
 &\cdot \int_{\mathbf{R}^n} e^{-|\xi|^2(1/4-it) + ix \cdot \xi} d\xi = \frac{1}{2^n} \frac{1}{\pi^{n/2}} \\
 &\cdot \int_{\mathbf{R}^n} e^{-(|\xi|^2(1/4-it) - ix \cdot \xi - |x|^2/(1-4it))} e^{-|x|^2/(1-4it)} d\xi = \frac{1}{2^n} \\
 &\cdot \frac{1}{\pi^{n/2}} e^{-|x|^2/(1-4it)} \int_{\mathbf{R}^n} e^{-|\xi \sqrt{1/4-it} + ix/(\sqrt{1-4it})|^2} d\xi.
 \end{aligned} \tag{33}$$

Now, we make the change of variables $\eta = \xi \sqrt{1/4-it} + ix/(\sqrt{1-4it})$ to get

$$\begin{aligned}
 u(t, x) &= \frac{1}{2^n} \frac{1}{\pi^{n/2}} e^{-|x|^2/(1-4it)} \int_{\mathbf{R}^n} e^{-|\eta|^2} \left(\frac{1}{4} - it\right)^{-n/2} d\eta \\
 &= \frac{1}{2^n} \frac{1}{\pi^{n/2}} e^{-|x|^2/(1-4it)} \left(\frac{1}{4} - it\right)^{-n/2} \pi^{n/2} \\
 &= (1 - 4it)^{-n/2} e^{-|x|^2/(1-4it)}.
 \end{aligned} \tag{34}$$

Hence, we obtain

$$u(t, x) = (1 - 4it)^{-n/2} e^{-|x|^2/(1-4it)}. \tag{35}$$

Now, we have to find a constant $c > 0$ such that, for every $t \in \mathbf{R}$, the function $p(t, x) = c^2 |u(t, x)|^2$ is a probability density function. The condition to be satisfied is the following:

$$\begin{aligned}
 1 &= \| |c|^2 u(t, x) \|_{L^2(\mathbf{R}^n)}^2 = \| |c|^2 u(0, x) \|_{L^2(\mathbf{R}^n)}^2 \\
 &= |c|^2 \int_{\mathbf{R}^n} e^{-2|x|^2} dx = |c|^2 \left(\frac{\pi}{2}\right)^{n/2},
 \end{aligned} \tag{36}$$

which implies $c = (\pi/2)^{-n/4}$. Therefore, the function

$$u(t, x) = \left(\frac{\pi}{2}\right)^{-n/4} (1 - 4it)^{-n/2} e^{-|x|^2/(1-4it)} \tag{37}$$

induces the probability density function:

$$\begin{aligned}
 p(t, x) &= |c|^2 |u(t, x)|^2 \\
 &= \left(\frac{\pi}{2}\right)^{-n/2} |1 - 4it|^{-n} \left| e^{-|x|^2/(1-4it)} \right|^2 \\
 &= \left(\frac{\pi}{2}\right)^{-n/2} |1 + 16t^2|^{-n/2} \left| e^{-(1+4it)|x|^2/(1+16t^2)} \right|^2 \\
 &= \left(\frac{\pi}{2}\right)^{-n/2} |1 + 16t^2|^{-n/2} e^{-2|x|^2/(1+16t^2)},
 \end{aligned} \tag{38}$$

so

$$p(t, x) = \left(\frac{\pi}{2}\right)^{-n/2} |1 + 16t^2|^{-n/2} e^{-2|x|^2/(1+16t^2)} \tag{39}$$

is going to be the generator of the family of distributions \mathcal{F} .

Remark 20. This procedure works because the Gaussian Distribution is, up to constants, a fixed point of the $1/\alpha$ -Characteristic Function. Indeed, if u is Gaussian, then $\phi_\alpha^u(e^{-|\cdot|^2})(\xi) = ce^{-|\xi|^2}$ for some normalizing constant $c > 0$. Moreover, the Schrödinger flow preserves Gaussian Distributions; namely, if your initial datum is Gaussian, then the solution is Gaussian for any future and past times.

2.4. Strichartz Estimates and Their Symmetries. In this subsection, we deduce the Strichartz Estimates for the Schrödinger equation in the case of probability distributions and discuss their symmetries. For clarity, we repeat here the definition of *admissible exponents* and the *Strichartz Estimate*.

Definition 21. Fix $n \geq 1$. One calls a set of exponents (q, r) admissible if $2 \leq q, r \leq +\infty$ and

$$\frac{2}{q} + \frac{n}{r} = \frac{n}{2}. \tag{40}$$

Theorem 22 (see [4-7]). For dimension $n = 1$ or $n = 2$ (and for any $n \geq 1$, supposing that Gaussians are maximizers) and (q, r) admissible pair, the Sharp Homogeneous Strichartz Constant $S_h(n, q, r) = S_h(n, r)$ defined by

$$\begin{aligned}
 S_h(n, r) &:= \sup \left\{ \frac{\|u\|_{L_t^q L_x^r(\mathbf{R} \times \mathbf{R}^n)}}{\|u\|_{L_x^2(\mathbf{R}^n)}} : u \in L_x^2(\mathbf{R}^n), u \neq 0 \right\}
 \end{aligned} \tag{41}$$

is given by

$$S_h(n, r) = 2^{n/4-n(r-2)/2r} r^{-n/2r}. \tag{42}$$

Moreover, if one defines $S_h(n) := S_h(n, 2 + 4/n, 2 + 4/n)$ by

$$S_h(n) = \sup \left\{ \frac{\|u\|_{L_{t,x}^{2+4/n}(\mathbf{R} \times \mathbf{R}^n)}}{\|u\|_{L_x^2(\mathbf{R}^n)}} : u \in L^2(\mathbf{R}^n), u \neq 0 \right\}, \tag{43}$$

then, for every $n \geq 1$ (always supposing that Gaussians are maximizers), one has that

$$S_h(n) = \left(\frac{1}{2} \left(1 + \frac{2}{n}\right)^{-n/2}\right)^{1/(2+4/n)}; \tag{44}$$

$S_h(n)$ is a decreasing function of n and

$$S_h(n) \rightarrow \frac{1}{(2e)^{1/2}}, \quad n \rightarrow +\infty. \tag{45}$$

For any $n \geq 1$ and (\bar{q}, \bar{r}) admissible pair, the Sharp Dual Homogeneous Strichartz Constant $S_d(q, r, n) = S_d(n, r)$ is defined by

$$\begin{aligned}
 S_d(n, r) &:= \sup \left\{ \frac{\left\| \int_{\mathbf{R}} e^{is\Delta} F(s) ds \right\|_{L_x^2}}{\|F\|_{L_t^{\bar{q}} L_x^{\bar{r}}}} : F \right. \\
 &\left. \in L_t^{\bar{q}} L_x^{\bar{r}}(\mathbf{R} \times \mathbf{R}^n), F \neq 0 \right\}.
 \end{aligned} \tag{46}$$

One has that $S_h(n, r) = S_d(n, r)$.

This is the version of the theorem on Strichartz Estimates without the restriction $\|u\|_{\mathbf{R}^n}^2 = 1$, as proved in [7]. From this, we can very easily deduce Theorem 6.

Proof of Theorem 6. Just substitute the condition $\|u\|_{\mathbf{R}^n}^2 = 1$ in all the statements. \square

As explained for example in [12], Strichartz Estimates are invariant by the following set of symmetries.

Lemma 23 (see [12]). *Let \mathcal{G} be the group of transformations generated by*

- (i) *space-time translations: $u(t, x) \mapsto u(t+t_0, x+x_0)$, with $t_0 \in \mathbf{R}, x_0 \in \mathbf{R}^n$;*
- (ii) *parabolic dilations: $u(t, x) \mapsto u(\lambda^2 t, \lambda x)$, with $\lambda > 0$;*
- (iii) *change of scale: $u(t, x) \mapsto \mu u(t, x)$, with $\mu > 0$;*
- (iv) *space rotations: $u(t, x) \mapsto u(t, Rx)$, with $R \in SO(n)$;*
- (v) *phase shifts: $u(t, x) \mapsto e^{i\theta} u(t, x)$, with $\theta \in \mathbf{R}$;*
- (vi) *Galilean transformations:*

$$u(t, x) \mapsto e^{(i/4)(|v|^2 t + 2v \cdot x)} u(t, x + tv), \tag{47}$$

with $v \in \mathbf{R}^n$.

Then, if u solves (27) and $g \in \mathcal{G}$, $v = g \circ u$ solves (27) also. Moreover, the constants $S_h(n, q, r)$, $S_d(n, q, r)$, and $S_s(n, q, r, \tilde{q}, \tilde{r})$ are left unchanged by the action of \mathcal{G} .

The only point here is that not all these symmetries leave the set of probability distributions $\mathcal{P}(\mathbf{R}^n)$ invariant. Therefore, we need to reduce the set of symmetries in our treatment and, in particular, we need to combine the scaling and the parabolic dilations in order to have all the family inside the space of probability distributions $\mathcal{P}(\mathbf{R}^n)$.

Lemma 24. *Consider $u_{\mu, \lambda} = \mu u(\lambda^2 t, \lambda x)$ such that $u(t, x) \in \mathcal{P}(\mathbf{R}^n)$ maximizes (6); then $\mu = \lambda^{n/2}$.*

Proof. Consider

$$\begin{aligned} 1 &= \|u_\lambda\|_{L^2(\mathbf{R}^n)}^2 = \mu^2 \int_{\mathbf{R}^n} |u(\lambda^2 t, \lambda x)|^2 dx \\ &= \mu^2 \lambda^{-n} \|u\|_{L^2(\mathbf{R}^n)}^2 = \mu^2 \lambda^{-n}, \end{aligned} \tag{48}$$

so $\mu = \lambda^{n/2}$. \square

Remark 25. We notice that some of the symmetries can be seen just at the level of the generator of the family u but not by the family of probability distributions $p_t(x)$. For example, the phase shifts $u(t, x) \mapsto e^{i\theta} u(t, x)$, with $\theta \in \mathbf{R}$, give rise to the same probability distribution function because $|e^{i\theta} u(t, x)|^2 = |u(t, x)|^2$ and, partially, the Galilean transformations $u(t, x) \mapsto e^{(i/4)(|v|^2 t + 2v \cdot x)} u(t, x + tv)$, with $v \in \mathbf{R}^n$, reduces to a space translation with $x_0 = vt$, since $|e^{(i/4)(|v|^2 t + 2v \cdot x)} u(t, x + tv)|^2 = |u(t, x + tv)|^2$. In some sense, the parameter θ can be seen as a latent variable.

Therefore, we have the complete set of probability distributions induced by the generator $u(t, x)$.

Theorem 26. *Consider $p_t(x) = |u(t, x)|^2$ a probability distribution function generated by $u(t, x)$ (see Section 2.3). Let \mathcal{S} be the group of transformations generated by*

- (i) *inertial-space translations and time translations: $p(t, x) \mapsto p(t + t_0, x + x_0 + vt)$, with $t_0 \in \mathbf{R}, x_0 \in \mathbf{R}^n$ and $v \in \mathbf{R}^n$;*
- (ii) *scaling-parabolic dilations: $u(t, x) \mapsto \lambda^n u(\lambda^2 t, \lambda x)$, with $\lambda > 0$;*
- (iii) *space rotations: $u(t, x) \mapsto u(t, Rx)$, with $R \in SO(n)$;*

Then, if u solves (27) and $g \in \mathcal{S}$, $v = g \circ u$ solves (27) also, $q_t(x) = |v(t, x)|^2$ is still a probability distribution for every $g \in \mathcal{S}$, and the constant $S_h(n, q, r)$ is left unchanged by the action of \mathcal{S} .

This theorem produces the following definition.

Definition 27. One calls Maximal Strichartz Family of Gaussian Distributions the following family of distributions:

$$\begin{aligned} \mathcal{F} := \left\{ p_t(x) = \left(\frac{\pi}{2}\right)^{-n/2} |R^T R|^{-1/2} |\lambda^2 + 16t^2|^{-n/2} \right. \\ \left. \cdot e^{-2(x-x_0-vt)^T (R^T R)^{-1} (x-x_0-vt) / (\lambda^2 + 16t^2)} : (t, \lambda) \right. \\ \left. \in \mathbf{R}^2, (x_0, v) \in \mathbf{R}^n \times \mathbf{R}^n, R \in SO(n) \right\}. \end{aligned} \tag{49}$$

Remark 28. Let $p(t, x)$ be the pdf defined in (39). Then, choose $\tilde{p}_t(x) \in \mathcal{F}$ with $R = Id$, $x_0 = v_0 = 0$, and $\lambda = 0$. This implies that $\tilde{p}_t(x) = p(t, x) \in \mathcal{F}$. For this reason, we call $p(t, x)$ the Family Generator of \mathcal{F} . We notice also that, in the definition of the family and with respect to Theorem 26, we used as scale parameter $\lambda^{1/2}$ instead of λ . This is done without loss of generality, since $\lambda > 0$.

Right away we can compute the *Variance-Covariance Matrix* and *Mean Vector* of the family.

Corollary 29. *Suppose \mathbf{X} is a random variable with pdf $p_t(x) \in \mathcal{F}$. Then its Expected Value is*

$$E[\mathbf{X}] := \mu = x_0 + vt \tag{50}$$

and its Variance is

$$\Sigma = \frac{1}{4} (\lambda^2 + 16t^2) (R^T R). \tag{51}$$

Proof. The proof is a direct computation. \square

Remark 30. We see here that, differently from the general family of Gaussian distributions, here the Mean Vector and the Variance-Covariance Matrix are related by a parameter, which represents the time flow.

3. The Fisher Information Metric of the Maximal Strichartz Family \mathcal{F}

Information geometry is a branch of mathematics that applies the techniques of differential geometry to the field of statistics and probability theory. This is done by interpreting probability distributions of a statistical model as the points of a Riemannian Manifold, forming in this way a statistical manifold. The Fisher Information Metric provides a natural Riemannian Metric for this manifold, but it is not the only possible one. With this tool, we can define and compute meaningful distances between probability distributions, in both the discrete and the continuous cases. Crucial is then the set of parameters on which a certain family of distributions is indexed and the geometrical structure of the parameter set is also crucial. We refer to [16] for a general reference on information geometry. The first one to introduce the notion of distance between two probability distributions has been Rao in [17], who used the Fisher Information Matrix as a Riemannian Metric on the space of parameters.

In this section, we restrict our attention to the Fisher Information Metric of the *Maximal Strichartz Family of Gaussian Distributions* \mathcal{F} and provide details on the additional structure that the family has with respect to the hyperbolic model of the general Family of Gaussian Distributions. See, for example, [18–20].

3.1. The Fisher Information Metric for the Multivariate Gaussian Distribution. First, we give the general definition of the Fisher Information Metric.

Definition 31. Consider a statistical manifold \mathcal{S} , with coordinates given by $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ and with probability density function $p(x; \theta)$. Here, x is a specific observation of the discrete or continuous random variables X . The probability is normalized, so that $\int_X p(x, \theta) dx = 1$ for every $\theta \in \mathcal{S}$. The Fisher Information Metric I_{ij} is defined by the following formula:

$$I_{ij}(\theta) = \int_X \frac{\partial \log p(x, \theta)}{\partial \theta_i} \frac{\partial \log p(x, \theta)}{\partial \theta_j} p(x, \theta) dx. \quad (52)$$

Remark 32. The integral is performed over all values x that the random variable X can take. Again, the variable θ is understood as a coordinate on the statistical manifold \mathcal{S} , intended as a Riemannian Manifold. Under certain regularity conditions (any that allows integration by parts), I_{ij} can be rewritten as

$$\begin{aligned} I_{ij}(\theta) &= - \int_X \frac{\partial^2 \log p(x, \theta)}{\partial \theta_i \partial \theta_j} p(x, \theta) dx \\ &= -E \left[\frac{\partial^2 \log p(x, \theta)}{\partial \theta_i \partial \theta_j} \right]. \end{aligned} \quad (53)$$

Now, to compute explicitly the *Fisher Information Matrix* of the family \mathcal{F} , we use the following theorem that you can find in [21].

Theorem 33. *The Fisher Information Matrix for an n -variate Gaussian Distribution can be computed in the following way. Let*

$$\mu(\theta) = [\mu_1(\theta), \mu_2(\theta), \dots, \mu_N(\theta)]^T \quad (54)$$

be the vector of Expected Values and let $\Sigma(\theta)$ be the Variance-Covariance Matrix. Then, the typical element $\mathcal{F}_{i,j}$, $0 \leq i, j < n$, of the Fisher Information Matrix for $X \sim \mathcal{N}(\mu(\theta), \Sigma(\theta))$ is

$$\mathcal{F}_{i,j} = \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} + \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right), \quad (55)$$

where $(\cdot)^T$ denotes the transpose of a vector, $\text{tr}(\cdot)$ denotes the trace of a square matrix, and

$$\begin{aligned} \frac{\partial \mu}{\partial \theta_i} &= \left[\frac{\partial \mu_1}{\partial \theta_i} \quad \frac{\partial \mu_2}{\partial \theta_i} \quad \dots \quad \frac{\partial \mu_n}{\partial \theta_i} \right]^T, \\ \frac{\partial \Sigma}{\partial \theta_i} &= \begin{bmatrix} \frac{\partial \Sigma_{1,1}}{\partial \theta_i} & \frac{\partial \Sigma_{1,2}}{\partial \theta_i} & \dots & \frac{\partial \Sigma_{1,n}}{\partial \theta_i} \\ \frac{\partial \Sigma_{2,1}}{\partial \theta_i} & \frac{\partial \Sigma_{2,2}}{\partial \theta_i} & \dots & \frac{\partial \Sigma_{2,n}}{\partial \theta_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \Sigma_{n,1}}{\partial \theta_i} & \frac{\partial \Sigma_{n,2}}{\partial \theta_i} & \dots & \frac{\partial \Sigma_{n,n}}{\partial \theta_i} \end{bmatrix}. \end{aligned} \quad (56)$$

Now, we have just to compute the *Fisher Information Matrix* entry by entry, following the theorem. We recall here that we are considering the following family of Gaussian Distributions:

$$\begin{aligned} \mathcal{F} := & \left\{ p(t, x) = \left(\frac{\pi}{2} \right)^{-n/2} |R^T R|^{-1/2} |\lambda^2 + 16t^2|^{-n/2} \right. \\ & \cdot e^{-2(x-x_0-vt)^T (R^T R)^{-1} (x-x_0-vt) / (\lambda^2 + 16t^2)} : (t, \lambda) \\ & \left. \in \mathbf{R}^2, (x_0, v) \in \mathbf{R}^n \times \mathbf{R}^n, R \in SO(n) \right\} \end{aligned} \quad (57)$$

and that, in particular, we have that the *Expected Value* of a random variable X with distribution belonging to the family \mathcal{F} is given by

$$\mu := x_0 + vt, \quad (58)$$

while the *Variance-Covariance Matrix* is given by

$$\Sigma := \frac{1}{4} (\lambda^2 + 16t^2) R^T R. \quad (59)$$

Remark 34. We remark again that μ and Σ depend on some common parameters, like the time t .

3.2. Proof of Theorem 8: The Spherical Multivariate Gaussian Distribution. Here, we consider the case in which $R^T R = \sigma^2 Id_{n \times n}$, namely, the case where the *Variance-Covariance Matrix* is given by $\Sigma := (1/4)(\lambda^2 + 16t^2)\sigma^2 Id_{n \times n}$. In this case, the vector of parameters θ is given by

$$\theta := (x_0^T, v_0^T, \lambda, \sigma^2, t)^T, \quad (60)$$

with x_0 and v_0 being $n \times 1$, while λ , σ^2 , and t are scalars. In order to fix the notation, we call $(\theta_1, \dots, \theta_n)^T := (x_0, \dots, x_n)$, $(\theta_{n+1}, \dots, \theta_{2n})^T := (v_0^0, \dots, v_0^n)$, $\theta_{2n+1} = \lambda$, $\theta_{2n+2} = \sigma^2$, and $\theta_{2n+3} = t$. Now, we want to compute all the coefficients of I_{ij} . We use the symmetry of the information matrix I_{ij} , so $I_{ij} = I_{ji}$. The relevant coefficients are

(i) $i = 1, \dots, n, j = 1, \dots, n$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, 1^i, \dots, 0)^T \frac{1}{\sigma^2} Id_{n \times n} (0, \dots, 1^j, \dots, 0) \quad (61) \\ &= \frac{1}{\sigma^2} \delta_{ij}; \end{aligned}$$

(ii) $i = 1, \dots, n, j = n + 1, \dots, 2n$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, 1^i, \dots, 0)^T \frac{1}{\sigma^2} Id_{n \times n} (0, \dots, t^j, \dots, 0) \quad (62) \\ &= \frac{t}{\sigma^2} \delta_{ij}; \end{aligned}$$

(iii) $i = 1, \dots, n, j = 2n + 1$

$$\mathcal{F}_{i,j} = 0, \quad (63)$$

because μ does not depend on λ and Σ does not depend on x_0 ;

(iv) $i = 1, \dots, n, j = 2n + 2$

$$\mathcal{F}_{i,j} = 0, \quad (64)$$

because μ does not depend on σ^2 and Σ does not depend on x_0 ;

(v) $i = 1, \dots, n, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, 1^i, \dots, 0)^T \frac{1}{\sigma^2} Id_{n \times n} (v_0^0, \dots, v_0^n) \quad (65) \\ &= \frac{1}{\sigma^2} v_0^i; \end{aligned}$$

(vi) $i = n + 1, \dots, 2n, j = n + 1, \dots, 2n$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, t^i, \dots, 0)^T \frac{1}{\sigma^2} Id_{n \times n} (0, \dots, t^j, \dots, 0) \quad (66) \\ &= \frac{t^2}{\sigma^2} \delta_{ij}; \end{aligned}$$

(vii) $i = n + 1, \dots, 2n, j = 2n + 1$

$$\mathcal{F}_{i,j} = 0, \quad (67)$$

because μ does not on λ and Σ does not depend on v_0 ;

(viii) $i = n + 1, \dots, 2n, j = 2n + 2$

$$\mathcal{F}_{i,j} = 0, \quad (68)$$

because μ does not depend on σ^2 and Σ does not depend on v_0 ;

(ix) $i = n + 1, \dots, 2n, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, 1^i, \dots, 0)^T \frac{1}{\sigma^2} Id_{n \times n} (v_0^0, \dots, v_0^n) \quad (69) \\ &= \frac{t}{\sigma^2} v_0^i; \end{aligned}$$

(x) $i = 2n + 1, j = 2n + 1$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma^2} Id \frac{2\lambda\sigma^2}{4} Id \frac{1}{\sigma^2} Id \frac{2\lambda\sigma^2}{4} Id \right) \quad (70) \\ &= \frac{\lambda^2}{8} \text{tr} (Id) = \frac{\lambda^2 n}{8}; \end{aligned}$$

(xi) $i = 2n + 1, j = 2n + 2$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma^2} Id \frac{2\lambda\sigma^2}{4} Id \frac{1}{\sigma^2} Id \frac{\lambda^2 + 16t^2}{4} Id \right) \quad (71) \\ &= n \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2); \end{aligned}$$

(xii) $i = 2n + 1, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma^2} Id \frac{2\lambda\sigma^2}{4} Id \frac{1}{\sigma^2} Id \frac{8t\sigma^2}{4} Id \right) = 2\lambda t n; \quad (72) \end{aligned}$$

(xiii) $i = 2n + 2, j = 2n + 2$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma^2} Id \frac{\lambda^2 + 16t^2}{4} Id \frac{1}{\sigma^2} Id \frac{\lambda^2 + 16t^2}{4} Id \right) \quad (73) \\ &= \frac{n}{32} \frac{(\lambda^2 + 16t^2)^2}{\sigma^4}; \end{aligned}$$

(xiv) $i = 2n + 2, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \operatorname{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \operatorname{tr} \left(\frac{1}{\sigma^2} Id \frac{\lambda^2 + 16t^2}{4} Id \frac{1}{\sigma^2} Id 8t\sigma^2 Id \right) \quad (74) \\ &= \frac{nt}{\sigma^2} (\lambda^2 + 16t^2); \end{aligned}$$

(xv) $i = 2n + 3, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} + \frac{1}{2} \operatorname{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= (v_0^0, \dots, v_0^n)^T \frac{1}{\sigma^2} Id_{n \times n} (v_0^0, \dots, v_0^n) \\ &\quad + \frac{1}{2} \operatorname{tr} \left(\frac{1}{\sigma^2} Id 8t\sigma^2 Id \frac{1}{\sigma^2} Id 8t\sigma^2 Id \right) \\ &= \frac{|v_0|^2}{\sigma^2} + 32nt^2. \end{aligned} \quad (75)$$

In conclusion, we have

$$I(\theta) = \begin{pmatrix} \frac{1}{\sigma^2} Id & \frac{t}{\sigma^2} Id & 0 & 0 & \frac{1}{\sigma^2} v_0 \\ \frac{t}{\sigma^2} Id & \frac{t^2}{\sigma^2} Id & 0 & 0 & \frac{t}{\sigma^2} v_0 \\ 0 & 0 & \frac{\lambda^2 n}{8} & n \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2) & 2\lambda tn \\ 0 & 0 & n \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2) & \frac{n}{32} \frac{(\lambda^2 + 16t^2)^2}{\sigma^4} & \frac{nt}{\sigma^2} (\lambda^2 + 16t^2) \\ \frac{1}{\sigma^2} v_0 & \frac{t}{\sigma^2} v_0 & 2\lambda tn & \frac{nt}{\sigma^2} (\lambda^2 + 16t^2) & \frac{|v_0|^2}{\sigma^2} + 32nt^2 \end{pmatrix}. \quad (76)$$

3.3. Proof of Theorem 8: The Elliptical Multivariate Gaussian Distribution. We define

$$\frac{1}{\sigma_i^2} Id := \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_n^2} \end{bmatrix}. \quad (77)$$

We define also

$$\begin{aligned} \frac{1}{\sigma^2} &:= \sum_{i=1}^n \frac{1}{\sigma_i^2}, \\ \frac{1}{\sigma^4} &:= \sum_{i=1}^n \frac{1}{\sigma_i^4}. \end{aligned} \quad (78)$$

Using this notations, we are going to compute the matrix I_{ij} . The relevant coefficients are

(i) $i = 1, \dots, n, j = 1, \dots, n$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, 1^i, \dots, 0)^T \frac{1}{\sigma_i^2} Id_{n \times n} (0, \dots, 1^j, \dots, 0) \quad (79) \\ &= \frac{1}{\sigma_i^2} \delta_{ij}; \end{aligned}$$

(ii) $i = 1, \dots, n, j = n + 1, \dots, 2n$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, 1^i, \dots, 0)^T \frac{1}{\sigma_i^2} Id_{n \times n} (0, \dots, t^j, \dots, 0) \quad (80) \\ &= \frac{t}{\sigma_i^2} \delta_{ij}; \end{aligned}$$

(iii) $i = 1, \dots, n, j = 2n + 1$

$$\mathcal{F}_{i,j} = 0, \quad (81)$$

because μ does not depend on λ and Σ does not depend on x_0 ;

(iv) $i = 1, \dots, n, j = 2n + 2$

$$\mathcal{F}_{i,j} = 0, \quad (82)$$

because μ does not depend on σ^2 and Σ does not depend on x_0 ;

(v) $i = 1, \dots, n, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, 1^i, \dots, 0)^T \frac{1}{\sigma^2} Id_{n \times n} (v_0^0, \dots, v_0^n) \quad (83) \\ &= \frac{1}{\sigma_i^2} v_0^i = \left(\frac{v_0^0}{\sigma_1^2}, \dots, \frac{v_0^n}{\sigma_n^2} \right); \end{aligned}$$

(vi) $i = n + 1, \dots, 2n, j = n + 1, \dots, 2n$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, t^i, \dots, 0)^T \frac{1}{\sigma^2} Id_{n \times n} (0, \dots, t^j, \dots, 0) \quad (84) \\ &= \frac{t^2}{\sigma_i^2} \delta_{ij}; \end{aligned}$$

(vii) $i = n + 1, \dots, 2n, j = 2n + 1$

$$\mathcal{F}_{i,j} = 0, \quad (85)$$

because μ does not depend on λ and Σ does not depend on v_0 ;

(viii) $i = n + 1, \dots, 2n, j = 2n + 2$

$$\mathcal{F}_{i,j} = 0, \quad (86)$$

because μ does not depend on σ^2 and Σ does not depend on v_0 ;

(ix) $i = n + 1, \dots, 2n, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} \\ &= (0, \dots, 1^i, \dots, 0)^T \frac{1}{\sigma^2} Id_{n \times n} (v_0^0, \dots, v_0^n) \quad (87) \\ &= \frac{t}{\sigma_i^2} v_0^i = t \left(\frac{v_0^0}{\sigma_1^2}, \dots, \frac{v_0^n}{\sigma_n^2} \right); \end{aligned}$$

(x) $i = 2n + 1, j = 2n + 1$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma^2} Id \frac{2\lambda\sigma^2}{4} Id \frac{1}{\sigma^2} Id \frac{2\lambda\sigma^2}{4} Id \right) \quad (88) \\ &= \frac{\lambda^2}{8} \text{tr}(Id) = \frac{\lambda^2 n}{8}; \end{aligned}$$

(xi) $i = 2n + 1, j = 2n + 2$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma_i^2} Id \frac{2\lambda\sigma_i^2}{4} Id \frac{1}{\sigma_i^2} Id \frac{\lambda^2 + 16t^2}{4} Id \right) \quad (89) \\ &= \sum_{i=1}^n \frac{\lambda}{16\sigma_i^2} (\lambda^2 + 16t^2) = \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2); \end{aligned}$$

(xii) $i = 2n + 1, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma_i^2} Id \frac{2\lambda\sigma_i^2}{4} Id \frac{1}{\sigma_i^2} Id \frac{8t\sigma_i^2}{4} Id \right) = 2\lambda t n; \end{aligned} \quad (90)$$

(xiii) $i = 2n + 2, j = 2n + 2$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma_i^2} Id \frac{\lambda^2 + 16t^2}{4} Id \frac{1}{\sigma_i^2} Id \frac{\lambda^2 + 16t^2}{4} Id \right) \\ &= \sum_{i=1}^n \frac{1}{32} \frac{((\lambda^2 + 16t^2)/4)^2}{\sigma_i^4} \\ &= \frac{1}{32} \frac{((\lambda^2 + 16t^2)/4)^2}{\sigma^4}; \end{aligned} \quad (91)$$

(xiv) $i = 2n + 2, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{tr} \left(\frac{1}{\sigma_i^2} Id \frac{\lambda^2 + 16t^2}{4} Id \frac{1}{\sigma_i^2} Id 8t\sigma^2 Id \right) \quad (92) \\ &= \sum_{i=1}^n \frac{t}{\sigma_i^2} (\lambda^2 + 16t^2) = \frac{t}{\sigma^2} (\lambda^2 + 16t^2); \end{aligned}$$

(xv) $i = 2n + 3, j = 2n + 3$

$$\begin{aligned} \mathcal{F}_{i,j} &= \frac{\partial \mu^T}{\partial \theta_i} \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} + \frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_j} \right) \\ &= (v_0^0, \dots, v_0^n)^T \frac{1}{\sigma_i^2} Id_{n \times n} (v_0^0, \dots, v_0^n) \\ &\quad + \frac{1}{2} \text{tr} \left(\frac{1}{\sigma_i^2} Id 8t\sigma_i^2 Id \frac{1}{\sigma_i^2} Id 8t\sigma_i^2 Id \right) \\ &= \sum_{i=1}^n \frac{|v_0^i|^2}{\sigma_i^2} + 32nt^2. \end{aligned} \quad (93)$$

In conclusion, we have

$$I(\theta) = \begin{pmatrix} \frac{1}{\sigma_i^2} Id & \frac{t}{\sigma_i^2} Id & 0 & 0 & \frac{1}{\sigma_i^2} v_0^j \\ \frac{t}{\sigma_i^2} Id & \frac{t^2}{\sigma_i^2} Id & 0 & 0 & \frac{t}{\sigma_i^2} v_0^j \\ 0 & 0 & \frac{\lambda^2 n}{8} & \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2) & 2\lambda tn \\ 0 & 0 & \frac{\lambda}{16\sigma^2} (\lambda^2 + 16t^2) & \frac{1}{32} \frac{(\lambda^2 + 16t^2)^2}{\sigma^4} & \frac{t}{\sigma^2} (\lambda^2 + 16t^2) \\ \frac{1}{\sigma_i^2} v_0^j & \frac{t}{\sigma_i^2} v_0^j & 2\lambda tn & \frac{t}{\sigma^2} (\lambda^2 + 16t^2) & \sum_{i=1}^n \frac{|v_0^j|^2}{\sigma_i^2} + 32nt^2 \end{pmatrix}. \tag{94}$$

This concludes the proof of Theorem 8.

3.4. The General Multivariate Gaussian Distribution. As pointed out in [18, 19], for general Multivariate Normal Distributions, the explicit form of the Fisher distance has not been computed in closed form yet even in the simple case where the parameters are $t = 0$, $\lambda = 0$, and $v_0 = 0$. From a technical point of view, as pointed out in [18, 19], the main difficulty arises from the fact that the sectional curvatures of the Riemannian Manifold induced by \mathcal{F} and endowed with the *Fisher Information Metric* are not all constant. We remark again here that the distance induced by our *Fisher Information Matrix* is centred at the *Maximal Strichartz Family of Gaussian Distributions*, to enlighten the difference between members of the *Maximal Strichartz Family of Gaussian Distributions* and other Gaussian Distributions, for which $R^T R = Id_{n \times n}$ is not necessarily satisfied. In particular, our metric distinguishes between Gaussians evolving through the PDE flow (see Section 2) and Gaussians who do not.

Remark 35. We say that two parameters α and β are orthogonal if the elements of the corresponding rows and columns of the *Fisher Information Matrix* are zero. Orthogonal parameters are easy to deal with in the sense that their maximum likelihood estimates are independent and can be calculated separately. In particular, for our family \mathcal{F} the parameters x_0 and v_0 are both orthogonal to both the parameters λ and σ^2 . Some partial results, for example, when either mean or variance is kept constant, can be deduced. See, for example, [18–20].

Remark 36. The Fisher Information Metric is not the only possible choice to compute distances between pdfs of the family of *Gaussian Distributions*. For example, in [20], the authors parametrize the family of normal distribution as the symmetric space $SL(n + 1)/SO(n + 1)$ endowed with the following metric:

$$ds^2 = \text{tr} \left(\Sigma^{-1} d\Sigma \Sigma^{-1} d\Sigma \right) - \frac{1}{n + 1} \left(\text{tr} \left(\Sigma^{-1} d\Sigma \right) \right)^2 + \frac{1}{2} d\mu^T \Sigma^{-1} d\mu. \tag{95}$$

Moreover, the authors in [20] computed the Riemann Curvature Tensor of the metric and, in any dimension, the distance between two normal distributions with the same mean and different variance and also the distance between two normal distributions with the same variance and different mean.

Remark 37. If we consider just the submanifold given by the restriction to the coordinates $i = 1, \dots, n$ and $i = 2n + 2$ on the ellipse $\lambda^2 + 16t^2 = 4$, we recover the hyperbolic distance:

$$ds_{\mathbb{H}}^2 := \frac{d\mu^2}{\sigma^2} + \frac{d\sigma^2}{2\sigma^4}. \tag{96}$$

The geometry, however, does not seem the one of a product space, at least considering the fact that mixed entries are not zero, in our parametrization.

4. Overdispersion, Equidispersion, and Underdispersion for the Family \mathcal{F}

As we said, Strichartz Estimates are a way to measure the dispersion caused by the flow of the PDE to which they are related. In statistics, dispersion explains how a distribution is spread-out. In this section, we connect the two closely related concepts (dispersion in statistics and PDEs) by introducing some measures of statistical dispersion like the *Index of Dispersion* in Definition 38 (see Section 4) which reflect the dispersive PDE nature of the *Maximal Strichartz Family of Gaussian Distributions*. We compute this *Index of Dispersion* for our family of distributions and show that it is consistent with PDE results.

Definition 38. Consider the norms $\|\cdot\|_a$ and $\|\cdot\|_b$ on the space of Variance-Covariance Matrices Σ and $\|\cdot\|_c$ on the space of mean values μ . One defines the following Index of Dispersion:

$$\mathcal{F}_M^{abc} := \|\Sigma(0)\|_a \times \frac{\|\Sigma(t)\|_b}{\|\mu(t)\|_c^4} \tag{97}$$

with $t \neq 0$ and where $\mu(t)$ is as follows:

$$\mu(t) := x_0 + vt, \tag{98}$$

while $\Sigma(t)$ is given by

$$\Sigma(t) := \frac{1}{4} (\lambda^2 + 16t^2) R^T R. \tag{99}$$

One calls \mathcal{F}_M^{abc} the *abc*-Dispersion Index of the Maximal Family of Gaussians and one calls

$$\mathcal{F}_M^a := \|\Sigma(0)\|_a \tag{100}$$

a-Static Dispersion Index of the Maximal Family of Gaussians. Moreover, one says that the distribution is

- (i) *abc*-overdispersed, if $\mathcal{F}_M^{abc} < 1$;
- (ii) *abc*-equidispersed, if $\mathcal{F}_M^{abc} = 1$;
- (iii) *abc*-underdispersed, if $\mathcal{F}_M^{abc} > 1$.

Analogously, one says that the distribution is

- (i) *a*-overdispersed, if $\mathcal{F}_M^a < 1$;
- (ii) *a*-equidispersed, if $\mathcal{F}_M^a = 1$;
- (iii) *a*-underdispersed, if $\mathcal{F}_M^a > 1$.

Here, we discuss some particular cases and compute the dispersion indexes \mathcal{F}_M^{abc} and \mathcal{F}_M^a for certain specific norms $\|\cdot\|_a, \|\cdot\|_b$ and $\|\cdot\|_c$.

(i) In the case $t = 0$, the *a*-Static Dispersion Index of the Maximal Family of Gaussians that we choose is given by the variance of the distribution. We choose $\|\Sigma\|_a = \det(\Sigma)$ and so we get

$$\mathcal{F}_M^a = \|\Sigma\|_a = \det(\Sigma) = \left(\frac{1}{4}\lambda^2\right)^n \det(R^T R). \tag{101}$$

Now, in the spherical case $R^T R = \sigma^2 Id$, one gets

$$\mathcal{F}_M^a = \left(\frac{1}{4}\lambda^2\sigma^2\right)^n. \tag{102}$$

So, the distribution is

- (i) *a*-overdispersed, if $(1/4)\lambda^2\sigma^2 < 1$;
- (ii) *a*-equidispersed, if $(1/4)\lambda^2\sigma^2 = 1$;
- (iii) *a*-underdispersed, if $(1/4)\lambda^2\sigma^2 > 1$.

Therefore, with $\|\Sigma\|_a = \det(\Sigma)$, the type of dispersion does not depend on the dimension n .

Remark 39. In the strictly Strichartz case $\sigma^2 = 1$, we have that the dispersion is measured just by the scaling factor λ .

Choosing instead $\|\Sigma\|_a = \text{tr}(\Sigma)$ as *a*-Static Dispersion Index of the Maximal Family of Gaussians, we have some small differences:

$$\mathcal{F}_M^a = \|\Sigma\|_a = \text{tr}(\Sigma) = \left(\frac{1}{4}\lambda^2\right) \text{tr}(R^T R). \tag{103}$$

Now, in the spherical case $R^T R = \sigma^2 Id$, we get

$$\mathcal{F}_M^a = n \left(\frac{1}{4}\lambda^2\sigma^2\right). \tag{104}$$

So, the distribution is

- (i) *a*-overdispersed, if $n(1/4)\lambda^2\sigma^2 < 1$;
- (ii) *a*-equidispersed, if $n(1/4)\lambda^2\sigma^2 = 1$;
- (iii) *a*-underdispersed, if $n(1/4)\lambda^2\sigma^2 > 1$.

So with $\|\Sigma\|_a = \text{tr}(\Sigma)$ the type of dispersion does depend on the dimension n .

(ii) In the case $t \in \mathbf{R}$, when t is different from zero, we can express Σ as a function of μ . In fact we have

$$t = \frac{(\mu - x_0) \cdot v_0}{|v_0|^2} \tag{105}$$

and so

$$\Sigma(\mu) = \frac{1}{4} \left(\lambda^2 + 16 \left(\frac{(\mu - x_0) \cdot v_0}{|v_0|^2} \right)^2 \right) R^T R. \tag{106}$$

For example, if now we choose $x_0 = 0$ and $\mu = v_0 = (1, 0, \dots, 0)^T$ we get

$$\Sigma(\mu) = \frac{1}{4} (\lambda^2 + 16) R^T R, \tag{107}$$

so for $\|\Sigma\|_a = \|\Sigma\|_b = \det(\Sigma)$ and $\|\mu\|_c^2 = \sum_{i=1}^n |\mu_i|^2$ we get, in the spherical case $R^T R = \sigma^2 Id$,

$$\mathcal{F}_M^{abc} = \left(\frac{1}{4}\lambda^2\sigma^2\right)^n \times \left(\frac{1}{4}(\lambda^2 + 16)\sigma^2\right)^n. \tag{108}$$

So the distribution is

- (i) *abc*-overdispersed, if $((1/4)\lambda^2\sigma^2)^n \times ((1/4)(\lambda^2 + 16)\sigma^2)^n < 1$;
- (ii) *abc*-equidispersed, if $((1/4)\lambda^2\sigma^2)^n \times ((1/4)(\lambda^2 + 16)\sigma^2)^n = 1$;
- (iii) *abc*-underdispersed, if $((1/4)\lambda^2\sigma^2)^n \times ((1/4)(\lambda^2 + 16)\sigma^2)^n > 1$.

Remark 40. In particular, from this, we notice that if at $t = 0$ the distribution is *a*-equidispersed, an instant after the distribution is *abc*-overdispersed, in fact

$$\begin{aligned} \mathcal{F}_M^{abc} &= \left(\frac{1}{4}\lambda^2\sigma^2\right)^n \times \left(\frac{1}{4}(\lambda^2 + 16)\sigma^2\right)^n \\ &= (1 + 4\sigma^2)^n > 1. \end{aligned} \tag{109}$$

This is in agreement with the dispersive properties of the family \mathcal{F} and legitimates, in some sense, our choice of Indexes of Dispersion. Moreover, if

$$\mathcal{F}_M^{abc} = (1 + 4\sigma^2)^n > 1 \tag{110}$$

is actually different from $\mathcal{F}_M^{abc} = 5^n$, namely, $\sigma^2 = 1$, we can argue that the Gaussian Distribution that we are analysing does not come from the Maximal Strichartz Family of Gaussian Distributions.

Remark 41. This index is different from the Fisher Index which is basically the variance to mean ratio

$$\mathcal{F}_F := \frac{\text{Var}(X)}{E[X]} \tag{111}$$

and it is the natural one for count data. The index \mathcal{F}_F is then more appropriate for families of distributions related to the Poisson distribution and that are dimensionless. In fact, in our case and in contrast with the Poisson case, we scale the Variance-Covariance Matrix as the square of the Expected Value: $\Sigma \approx \mu^2$.

Remark 42. The characterization of the Gaussian Distribution given by Theorems 6 and 22 can be used also to give a measure of dispersion with respect to the Maximal Family of Gaussian Distributions, considering the Strichartz Norm:

$$\mathcal{F}_S := \frac{\|e^{it\Delta}u_0(x)\|_{L^r_t L^r_x}}{S(n,r)}. \tag{112}$$

By Theorem 6, one has that $0 \leq \mathcal{F}_S \leq 1$. When the index is close to one, the distribution is close, in some sense, to the family \mathcal{F} , while, when the index is close to zero, the distribution is very far from \mathcal{F} . This index clearly does not distinguish between distributions in the family \mathcal{F} . It would be very interesting to see if the closeness to one of the Indexes of Dispersion \mathcal{F}_S computed on a general distribution implies a proximity to the Maximal Family of Gaussian Distributions from the distribution point of view also and not just from the point of view of the dispersion.

5. Partial Stochastic Ordering on \mathcal{F}

Using the concept of *Index of Dispersion*, we can give a *Partial Stochastic Order* to the family \mathcal{F} . For a more complete treatment on *Stochastic Orders*, we refer to [14]. We start the analysis of this section with the definition of *Mean-Preserving Spread*.

Definition 43. A Mean-Preserving Spread (MPS) is a map from $\mathcal{P}(\mathbf{R}^n)$ to itself

$$p(x; \theta_1) \longrightarrow p(x; \theta_2), \tag{113}$$

where $p(x; \theta_1)$ and $p(x; \theta_2)$ are, respectively, the pdf of the random variables X_1 and X_2 with the property of leaving the Expected Value unchanged:

$$\mu_{X_1}(\theta_1) = \mu_{X_2}(\theta_2), \tag{114}$$

for any θ_1 and θ_2 in the space of parameters.

The concept of a *Mean-Preserving Spread* provides a partial ordering of probability distributions according to their level of dispersion. We then give the following definition.

Definition 44. Consider two random variables X_1 and X_2 such that $\mu_{X_1}(\theta_1) = \mu_{X_2}(\theta)$, for any θ_1 and θ_2 . One says that the two random variables are ordered according to their

Dispersion Index \mathcal{F} if and only if the following condition is satisfied:

$$\begin{aligned} X_1 < X_2 &\iff \\ \mathcal{F}(X_2) &\leq \mathcal{F}(X_1). \end{aligned} \tag{115}$$

Now, we give some examples of ordering according to the Indexes of Dispersion that we discussed previously.

(i) In case $t = 0$, we choose $\|\Sigma\|_a = \det(\Sigma)$ and so we get

$$\mathcal{F}_M^a = \|\Sigma\|_a = \det(\Sigma) = \left(\frac{1}{4}\lambda^2\right)^n \det(R^T R). \tag{116}$$

Now, in the spherical case $R^T R = \sigma^2 Id$, one gets

$$\mathcal{F}_M^a = \left(\frac{1}{4}\lambda^2\sigma^2\right)^n. \tag{117}$$

Using this index, we have the following partial ordering:

$$\begin{aligned} X_1 < X_2 &\iff \\ \lambda_2^2\sigma_2^2 &\leq \lambda_1^2\sigma_1^2. \end{aligned} \tag{118}$$

This order does not depend on the dimension n . By choosing instead $\|\Sigma\|_a = \text{tr}(\Sigma)$, we obtain

$$\mathcal{F}_M^a = \|\Sigma\|_a = \text{tr}(\Sigma) = \left(\frac{1}{4}\lambda^2\right) \text{tr}(R^T R). \tag{119}$$

Now, again in the spherical case $R^T R = \sigma^2 Id$, one gets

$$\begin{aligned} X_1 < X_2 &\iff \\ \lambda_2^2\sigma_2^2 &\leq \lambda_1^2\sigma_1^2, \end{aligned} \tag{120}$$

which is the same ordering as before. This order does not depend on the dimension n again and this seems to suggest that even if the value of the *Dispersion Index* might depend on the choice of the norms, the *Partial Order* is less sensible to it.

Remark 45. In the strictly Strichartz case $\sigma^2 = 1$, we have that the Stochastic Order is given just by the scaling factor λ .

(ii) In the case when t is different from zero, we have

$$\Sigma(\mu) = \frac{1}{4} \left(\lambda^2 + 16 \left(\frac{(\mu - x_0) \cdot v_0}{|v_0|^2} \right)^2 \right) R^T R. \tag{121}$$

If now we choose $x_0 = 0$ and $\mu = v_0 = (1, 0, \dots, 0)^T$, we get

$$\Sigma(\mu) = \frac{1}{4} (\lambda^2 + 16) R^T R, \tag{122}$$

so, for $\|\Sigma\|_a = \|\Sigma\|_b = \det(\Sigma)$ and $\|\mu\|_c^2 = \sum_{i=1}^n |\mu_i|^2$, we get, in the spherical case $R^T R = \sigma^2 Id$, the following *Partial Order*:

$$\begin{aligned} X_1 < X_2 &\iff \left(\frac{1}{4}\lambda_2^2\sigma_2^2\right)^n \times \left(\frac{1}{4}(\lambda_2^2 + 16)\sigma_2^2\right)^n \\ &\leq \left(\frac{1}{4}\lambda_1^2\sigma_1^2\right)^n \times \left(\frac{1}{4}(\lambda_1^2 + 16)\sigma_1^2\right)^n. \end{aligned} \tag{123}$$

Remark 46. Again, in the strictly Strichartz case $\sigma^2 = 1$, we have that the Stochastic Order is given just by the scaling factor λ .

Remark 47. In the case of the the a -Static Dispersion Index of the Maximal Family of Gaussians \mathcal{F}_M^a , the role of σ^2 and λ^2 seems interchangeable. This suggests a dimensional reduction in the parameter space, but, when $t \neq 0$, σ^2 and the parameter λ decouple and start to play a slightly different role. This suggests again a way to distinguish between Gaussian Distributions which come from the family \mathcal{F} and Gaussians which do not and so to distinguish between Gaussians which are solutions of the Linear Schrödinger Equation and Gaussians which are not.

Remark 48. Using the definition of Entropy, we deduce that, for Gaussian Distributions, $h(X) = (1/2) \log(2\pi e)^n \det(\Sigma)$. We see that, for our family \mathcal{F} , the Entropy increases, every time we increase λ, σ^2 , and t , but not when we increase x_0 and v_0 . In particular, the fact that the Entropy increases with t is in accordance with the Second Principle of Thermodynamics.

Remark 49. It seems that the construction of similar indexes can be performed in more general situations. In particular, we think that an index similar to \mathcal{F}_M^{abc} can be computed in every situation in which a family of distributions has the Variance-Covariance Matrix and the Expected Value which depend on common parameters.

6. Conclusions

In this paper, we have constructed and studied the *Maximal Strichartz Family of Gaussian Distributions*. This subfamily of the family of Gaussian Distributions arises naturally in the context of Partial Differential Equations and Harmonic Analysis, as the set of maximizers of certain functionals introduced by Strichartz [4] in the context of the *Schrödinger Equation*. We analysed the *Fisher Information Matrix* of the family and we showed that this matrix possesses an extrastructure with respect to the general family of Gaussian Distributions. We studied the spherical and elliptical case and computed explicitly the *Fisher Information Metric* in both cases. We interpreted the *Fisher Information Metric* as a distance which can distinguish between Gaussians which maximize the *Strichartz Norm* and Gaussians which do not and also as a distance between Gaussians which are solutions of the *Linear Schrödinger Equation* and Gaussians which are not. After this, we introduced some measures of statistical dispersion that we called *abc-Dispersion Index of the Maximal Family of Gaussian Distributions* and *a-Static Dispersion Index of the Maximal Family of Gaussian Distributions*. We showed that these Indexes of Dispersion are consistent with the dispersive nature of the Schrödinger Equation and can be used again to distinguish between Gaussians belonging to the family \mathcal{F} and other Gaussians. Moreover, we showed that our Indexes of Dispersion induce a *Partial Stochastic Order* on the *Maximal Strichartz Family of Gaussian Distributions*, which is in accordance with the flow of the PDE.

Appendix

In this Appendix, we give the proof of Theorem 3 (for completeness) and we will use the concept of $1/\alpha$ -Characteristic Function to define a generalized type of Moments that exist also for the *Multivariate Cauchy Distribution*.

A. Proof of Theorem 3

The proof is very simple and can be found in several places (see, e.g., a nice treatment in [22]). We report here, for completeness, the computation in the case $n = 1$. We consider the variational derivative of $h(X)$ with the constrain of f being a probability distribution and with the constraint of having a fixed variance σ^2 . This gives rise to the following *Euler Lagrange Equation* with two *Lagrangian multipliers* λ_0 and λ_2 :

$$\mathcal{L}(v) = \int_{\mathbf{R}} v(x) \ln(v(x)) dx - \lambda_0 \left(1 - \int_{\mathbf{R}} v(x) dx \right) - \lambda_2 \left(\sigma^2 - \int_{\mathbf{R}} v(x) (x - \mu)^2 dx \right), \tag{A.1}$$

where $v(x)$ is some function with *Expected Value* μ . The two *Lagrangian multipliers* appear because of the two constraints. One constraint is related to the normalization condition

$$\int_{\mathbf{R}} v(x) dx = 1 \tag{A.2}$$

and the other is related to the requirement of fixed variance:

$$\sigma^2 = \int_{\mathbf{R}} v(x) (x - \mu)^2 dx. \tag{A.3}$$

Now, we take the variational derivative of the functional \mathcal{L} . To be at a critical point, we need to impose that this variational derivative is zero. Therefore, we get

$$0 = L'(v) = \frac{d}{dt} \Big|_{t=0} L(v(x) + tg(x)) = \int_{-\infty}^{\infty} g(x) (\ln(v(x)) + 1 + \lambda_0 + \lambda(x - \mu)^2) dx. \tag{A.4}$$

Since this must hold for any variation $g(x)$, the term in brackets must be zero, and so, solving for $v(x)$ yields

$$v(x) = e^{-\lambda_0 - 1 - \lambda(x - \mu)^2}. \tag{A.5}$$

Now, we use the constraint of the problem and solve for λ_0 and λ_2 . From $\int_{\mathbf{R}} v(x) dx = 1$, we get the condition

$$\lambda_2^{-1/2} \pi^{1/2} = e^{\lambda_0 + 1} \tag{A.6}$$

and from $\sigma^2 = \int_{\mathbf{R}} v(x) (x - \mu)^2 dx$, we get

$$\sigma^2 e^{\lambda_0 + 1} \lambda_2^{3/2} = \frac{\pi^{1/2}}{2}. \tag{A.7}$$

Solving for λ_0 and λ_2 we get $\lambda_0 = (1/2) \log(2\pi\sigma^2) - 1$ and $\lambda_2 = 1/2\sigma^2$ which altogether give the *Gaussian Distribution*:

$$v(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}. \tag{A.8}$$

B. On the $1/\alpha$ -Momenta of Order k and the Cauchy Distribution

In this subsection, we discuss another application of the concept of $1/\alpha$ -Characteristic Functions. In particular, we build $1/\alpha$ -Momenta in a similar way of what happens for the usual characteristic function and usual Momenta. We apply this tool to the case of the Cauchy Distribution and see that, in certain cases, in contrast to the well-known case of $\alpha = 1$, we can build some finite generalized Momenta. We refer to [15] for a more detailed discussion on $1/\alpha$ -Characteristic Functions.

Definition 50. Consider $\phi_\alpha^u(\xi)$ the $1/\alpha$ -Characteristic Function of u :

$$\phi_\alpha^u(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix \cdot \xi} u(x) dx. \tag{B.1}$$

Then, if $\phi_\alpha^u(\xi)$ is k times that is continuously differentiable on \mathbb{R} , one defines the $1/\alpha$ -Moment of order k by the formula

$$E^{1/\alpha} [X^k] := (-i)^k \phi_\alpha^u(\xi)(0). \tag{B.2}$$

Proof. This is a direct and simple computation. □

Remark 51. Here, we do not consider the possibility of different roots of the unity that can appear in the computation of the $1/\alpha$ -Characteristic Function. We refer, for the precise theory, to [15].

From now on, we concentrate only on the case of the Multivariate Cauchy Distribution.

Definition 52. One says that a random variable X is distributed as a Multivariate Cauchy Distribution if and only if its pdf $f(\mathbf{x}; \mu, \Sigma, n)$ takes the form

$$f(\mathbf{x}; \mu, \Sigma, n) = \frac{\Gamma((1+n)/2)}{\Gamma(1/2) \pi^{n/2} |\Sigma|^{1/2} [1 + (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)]^{(1+n)/2}} \tag{B.3}$$

Now, we want to determine for which $\alpha > 0$ and $k, n = 1, 2, \dots$ the $1/\alpha$ -Momenta of order k exist and are finite. In other words, we want to find for which values of n, k , and α , we have that $|E^{1/\alpha}[X^k]| < +\infty$. Therefore, we compute

$$\begin{aligned} |E^{1/\alpha} [X^k]| &\leq C \int_{\mathbb{R}^n} \frac{|x|^k}{(1 + |x|^2)^{(1+n)/2\alpha}} dx \\ &\leq C \int_1^{+\infty} \frac{\rho^{k+n-1}}{\rho^{(n+1)/\alpha}} d\rho \\ &\leq C \int_1^{+\infty} \frac{1}{\rho^{(n+1)/\alpha - n - k + 1}}, \end{aligned} \tag{B.4}$$

where the constant C may vary from step to step. So, we have that the $1/\alpha$ -Momenta of order k exist and are finite, when

$$\begin{aligned} |E^{1/\alpha} [X^k]| < +\infty &\iff \\ \frac{n+1}{\alpha} - n - k + 1 > 1, \end{aligned} \tag{B.5}$$

namely, if and only if the order of the momentum k satisfies the following condition:

$$k < \frac{n+1}{\alpha} - n. \tag{B.6}$$

In the case $n = k = 1$, we need $0 < \alpha < 1$ and, in general, we need $0 < \alpha < (n+1)/(n+k)$ in order to have that the $1/\alpha$ -Moment of order k , $E^{1/\alpha}[X^k]$ is well defined.

Competing Interests

The author declares that he has no competing interests.

Acknowledgments

The author wants to thank their family and Victoria Ban for their constant support. They thank Professor Maung Min-Oo for useful discussions on the Fisher Information Metric. The author thanks also their Supervisor Professor Narayanaswamy Balakrishnan for his constant help and guidance and for an inspiring talk on dispersion indexes.

References

- [1] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, New York, NY, USA, 1982.
- [2] S. Kotz, N. Balakrishnan, and N. L. Johnson, *Continuous Multivariate Distributions, Vol. 1: Models and Applications*, Wiley, New York, NY, USA, 2nd edition, 2000.
- [3] T. Tao, *Nonlinear Dispersive Equations: Local and Global Analysis*, BMS Regional Conference Series in Mathematics, No. 106, 2006.
- [4] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations," *Duke Mathematical Journal*, vol. 44, no. 3, pp. 705–714, 1977.
- [5] M. Keel and T. Tao, "Endpoint Strichartz estimates," *American Journal of Mathematics*, vol. 120, no. 5, pp. 955–980, 1998.
- [6] M. Christ and R. Quilodr an, "Gaussians rarely extremize adjoint Fourier restriction inequalities for paraboloids," *Proceedings of the American Mathematical Society*, vol. 142, no. 3, pp. 887–896, 2014.
- [7] A. Selvitella, "Remarks on the sharp constant for the Schrödinger Strichartz estimate and applications," *Electronic Journal of Differential Equations*, vol. 2015, no. 270, pp. 1–19, 2015.
- [8] J. Bennett, N. Bez, C. Jeavons, and N. Pattakos, "On sharp bilinear Strichartz estimates of Ozawa-Tsutsumi type," *Journal of the Mathematical Society of Japan*. *In Press*, <http://mathsoc.jp/publication/JMSJ/pdf/JMSJ7238.pdf>.
- [9] D. Hundertmark and V. Zharnitsky, "On sharp Strichartz inequalities in low dimensions," *International Mathematics Research Notices*, vol. 2006, Article ID 34080, 18 pages, 2006.
- [10] J. Jiang and S. Shao, "On characterization of the sharp Strichartz inequality for the Schrödinger equation," *Analysis & PDE*, vol. 9, no. 2, pp. 353–361, 2016.
- [11] S. Klainerman and M. Machedon, "Remark on Strichartz-type inequalities," *International Mathematics Research Notices*, no. 5, pp. 201–220, 1996.
- [12] D. Foschi, "Maximizers for the Strichartz inequality," *Journal of the European Mathematical Society*, vol. 9, no. 4, pp. 739–774, 2007.

- [13] M. Kunze, "On the existence of a maximizer for the Strichartz inequality," *Communications in Mathematical Physics*, vol. 243, no. 1, pp. 137–162, 2003.
- [14] E. L. Lehmann, "Ordered families of distributions," *The Annals of Mathematical Statistics*, vol. 26, pp. 399–419, 1955.
- [15] A. Selvitella, "On $1/\alpha$ -characteristic functions and applications to asymptotic statistical inference," *Communications in Statistics—Theory and Methods*, 2016.
- [16] S. Amari and H. Nagaoka, *Methods of Information Geometry*, vol. 191 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI, USA, 2000.
- [17] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," *Bulletin of the Calcutta Mathematical Society*, vol. 37, pp. 81–91, 1945.
- [18] S. I. R. Costa, S. A. Santos, and J. E. Strapasson, "Fisher information distance: a geometrical reading," *Discrete Applied Mathematics*, vol. 197, pp. 59–69, 2015.
- [19] S. I. R. Costa, S. A. Santos, and J. E. Strapasson, "Fisher information matrix and hyperbolic geometry," in *Proceedings of the IEEE Information Theory Workshop (ITW '05)*, pp. 34–36, IEEE, August-September 2005.
- [20] M. Lovrić, M. Min-Oo, and E. A. Ruh, "Multivariate normal distributions parametrized as a Riemannian symmetric space," *Journal of Multivariate Analysis*, vol. 74, no. 1, pp. 36–48, 2000.
- [21] K. V. Mardia and R. J. Marshall, "Maximum likelihood estimation of models for residual covariance in spatial regression," *Biometrika*, vol. 71, no. 1, pp. 135–146, 1984.
- [22] K. Conrad, "Probability Distributions and Maximum Entropy," <http://www.math.uconn.edu/~kconrad/blurbs/analysis/entropypost.pdf>.